

Weekly Report

梅鸿辉

June 11, 2018

1. Vis论文投稿总结

review中最主要的一点：**Evaluation**不够完善，包括缺乏对已有类似系统的对比，以及对实验数据、试验结果的详细解释。

其他方面：

- 缺乏针对performance-accuracy trade-off的详细探讨
- 开源的demo

主要修改方案：

- 重构代码，使得可以有更高的效率，可以和其他工具对比，并给出开源项目
- 重新设计试验
- 修订论文

时间安排见文末表格

2. CHI投稿项目idea

VisEvo暂时没有太好的想法，目前主要在考虑陈老师提出的基于预测的可视推荐，基本思路：

- 通过预测算法预测未来短时间内数据可能的变化，预测算法有轨迹预测/LSTM（加强的RNN）/回归等
- 将视图的价值评估拓展到一个时间段上的整体评估
- 通过一个简单的调度算法，或者混合一些决策算法，进行安排，排除每个时刻推荐的top-K
- 相应的可视化表达

目前我觉得可能LSTM（RNN）会更好。作为拥有记忆功能的神经网络，RNN可以用作预测、决策等，可能可以将预测和推荐统一一并执行。

另一方面，如果采用决策算法来进行计算，可选的包括简单的赌博机问题中常见算法（也是SeeDB采用的方法）、或者使用基于蒙特卡洛随机的树搜索（与下棋的蒙特卡洛树可能有些实现上的差异）将未来的不确定性纳入考量。后者比较依赖于预测算法的具体结果，比如是否会给出多个待选结果，是否能给出置信度信息（没有的话可能可以直接假设一个预测周围的高斯分布，方差由与测试数据的对比计算得到）。

Papaer Reading

2.1 可视推荐系统

[1] A. Anand and J. Talbot, “Automatic Selection of Partitioning Variables for Small Multiple Displays,” *IEEE Trans. Vis. Comput. Graph.*, vol. 22, no. 1, pp. 669–677, 2016.

通过一些指标（Scagnostics）对数据进行划分，生成各种截面视图

[2] J. Seo and B. Shneiderman, “A Rank-by-Feature Framework for Interactive Exploration of Multidimensional Data,” *Inf. Vis.*, vol. 4, no. 2, pp. 96–113, 2005.

[3] S. Johansson and J. Johansson, “Interactive Dimensionality Reduction Through User-defined Combinations of Quality Metrics,” *Computer (Long. Beach. Calif.)*, vol. 15, no. 6, pp. 993–1000, 2009.

[4] E. Bertini, A. Tatu, and D. Keim, “Quality metrics in high-dimensional data visualization: An overview and systematization,” *IEEE Trans. Vis. Comput. Graph.*, vol. 17, no. 12, pp. 2203–2212, 2011.

看看其中一些指标

[5] G. Koutrika, “Modern Recommender Systems : from Computing Matrices to Thinking with Neurons,” pp. 1651–1654, 2018.

可视推荐系统综述：包括基于价值评估函数的、就与赌博机问题（决策问题）的和基于神经网络的。

2.2 CHI'18

[1] Z. Liu et al., “Data Illustrator: Augmenting Vector Design Tools with Lazy Data Binding for Expressive Visualization Authoring,” *Proc. 2018 CHI Conf. Hum. Factors Comput. Syst. - CHI ' 18*, pp. 1–13, 2018.

[2] H. Xia, N. Henry Riche, F. Chevalier, B. De Araujo, and D. Wigdor, “DataInk: Direct and Creative Data-Oriented Drawing,” *Proc. 2018 CHI Conf. Hum. Factors Comput. Syst. - CHI ' 18*, pp. 1–13, 2018.

[3] Y. Wang et al., “InfoNice : Easy Creation of Information Graphics,” pp. 1–12, 2018.

三篇交互式可视化构建工具的文章，作为VisEvo投CHI的参考。其中第一篇是Best paper(?)，对于如何让可视化绘制工具可以给设计师使用写的很详细

2.3 其他

[1] J. Zhou, A. K. H. Tung, W. Wu, and W. S. Ng, “A ‘semi-lazy’ approach to probabilistic path prediction in dynamic environments,” KDD ’ 13 Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discov. data Min., p. 748, 2013.

轨迹预测

[2] T. Kraska, A. Beutel, E. H. Chi, and J. Dean, “The Case for Learned Index Structures,” no. 1, pp. 1–27.

通过深度学习模拟B-tree: 对给定的数据集, 机器学习模型不存在过拟合的缺陷

[3] M. A. Borkin et al., “What makes a data visualization memorable?,” IEEE Trans. Vis. Comput. Graph., vol. 19, no. 12, pp. 2306–2315, 2013.

哪种可视化更容易被记忆? 两个结论: 1, 可视化比常规图片更难记忆, 接近人脸照片; 2, 出乎意料的, 复杂的图(矩阵、树、网络等)比通用图表(柱状图、折线图等)更容易被记忆

计划-已完成

TASK	DESCRIPTION	NOTE
大黑书修订	初稿	可能需要进一步修订

计划-已完成

TASK	DESCRIPTION	SCHEDULE
大黑书修订	继续修改	
CHI的idea		下周继续讨论
专利(两篇)	初稿完成(1/2)	与律师沟通

计划-中期

TASK	DESCRIPTION	SCHEDULE
RSATree代 码 重构	后端C++化	暂时可以开始一部分，视CHI的idea构思进展而定
RSATree论文	修订后投TVCG	CHI之后开始

计划-长期

TASK	DESCRIPTION	SCHEDULE
毕业论文	目前定位为可是设计方向	开始考虑一下整体构思

Works Progresses

TASK	PROGRESS	TODO	ISSUES	DATE
RSATree	修订投TVCG	整理代码		
RSATree专利		与律师沟通		下周
大黑书修订				
VisEvo		idea evaluation		
电子学报	已进入最后阶段			
ECharts论文	完成proof	等待最终发布		
分辨率自适应 可视化		学习/咨询相关理论基础		